

Checking File Integrity

Category: File Transfers

DRAFT

This article is being reviewed for completeness and technical accuracy.

It is a good practice to check that your data are complete and accurate before and after a file transfer. A common way for checking data integrity is to compute a checksum of the data.

There are multiple algorithms and programs that one can use for computing a checksum. A good checksum algorithm will yield a different result with high probability when the data is accidentally corrupted. If the checksums obtained before and after the transfer match, the data is almost certainly not corrupted.

On NAS HECC systems, the following programs are available:

- *sum*

computes a checksum using BSD sum or System V sum algorithm; also counts the number of blocks (1KB-block or 512B-block) in a file

- *cksum*

computes a cyclic redundancy check (CRC) checksum; also counts the number of bytes in a file

- *md5sum*

computes a 128-bit MD5 checksum which is represented by a 32-character hexadecimal number

For example,

```
%ls -l foo
-rw----- 1 username groupid 67358 Nov 15 11:49 foo

%sum foo
50063      66

%cksum foo
269056887 67358 foo

%md5sum foo
```

```
cfe0fc62607e9dc6ea0c231982316b75  foo
```

md5sum is more reliable than *sum* or *cksum* for detecting accidental file corruption, as the chances of accidentally having two files with identical MD5 checksum are extremely small. It is installed by default in most Unix, Linux, and Unix-like operating systems. Users are recommended to compute the *md5sum* of a file before and after the transfer.

The following example shows that the file *foo* is complete and accurate after the transfer based on its md5sum.

```
pfe1% md5sum foo
cfe0fc62607e9dc6ea0c231982316b75  foo

pfe1% scp foo local_username@your_localhost:

your_localhost%md5sum foo
cfe0fc62607e9dc6ea0c231982316b75  foo
```

See **sum**, **cksum**, **md5sum** man pages for more information.

Article ID: 243

Last updated: 16 May, 2011

Data Storage & Transfer -> File Transfers -> Checking File Integrity

<http://www.nas.nasa.gov/hecc/support/kb/entry/243/?ajax=1>